A. Notation

$\mathcal{R}(A)$	range space of the operator A
$\mathcal{N}(A)$	null space of the operator A
A'	the transpose of the matrix A
A^*	the adjoint of the operator A, or the complex-conjugate-transpose of the matrix A
$A > 0$	a positive-definite matrix
$A \geq 0$	a positive-semi-definite matrix
$\lambda_i(A)$	ith eigenvalue of A
$\text{Spec}(A)$	the set of eigenvalues of A
$\rho(A)$	spectral radius of A $\quad = \max_i \lvert\lambda_i(A)\rvert$
$\sigma_i(A)$	ith singular value of A (in descending order)
$\overline{\sigma}$	largest singular value
$\underline{\sigma}$	smallest nonzero singular value

B. Linear Operators

1 Let $\mathbf{V}$ and $\mathbf{W}$ be vector spaces over the same base field $\mathbf{F}$.

Definition A *linear operator* is a mapping

$$\mathcal{M} : \mathbf{V} \longrightarrow \mathbf{W}$$

such that for all $v_1, v_2 \in \mathbf{V}$ and all $\alpha \in \mathbf{F}$

(a) $\mathcal{M}(v_1 + v_2) = \mathcal{M}(v_1) + \mathcal{M}(v_2)$ (additivity)

(b) $\mathcal{M}(\alpha v_1) = \alpha \mathcal{M}(v_1)$ (homogeneity)

2 **Examples** The following operators are linear:

- ◇ $\mathcal{M} : \mathbf{C}^n \to \mathbf{C}^m : v \to Av$ where $A \in \mathbf{C}^{m\times n}$
- ◇ $\mathcal{M} : \mathcal{C}(-\infty, \infty) \to \mathbf{R} : f(t) \to f(0)$.
  Is the operator $\mathcal{M}$ above familiar ?
- ◇ Suppose $h(t) \in L_2[a, b]$ and consider

  $$\mathcal{M} : L_2[a, b] \to \mathbf{R} : f(t) \to \int_a^b h(t)f(t)dt$$

  Why do we insist that $f \in L_2[a, b]$ above ?
- ◇ Let $A, B, X \in \mathbf{R}^{n\times n}$ and consider

  $$\mathcal{M} : \mathbf{R}^{n\times n} \to \mathbf{R}^{n\times n} : X \to AX + XB$$

3 Linear operators on finite dimensional vector spaces are matrices with the action of matrix-vector multiplication.

Theorem *Consider a linear operator $\mathcal{M} : \mathbf{V} \to \mathbf{W}$ where $dim(\mathbf{V}) = n, dim(\mathbf{W}) = m$. Let $B = \{b_1, \cdots, b_n\}$ and $C = \{c_1, \cdots, c_n\}$ be basis for $\mathbf{V}$ and $\mathbf{W}$ respectively. Suppose $\mathcal{M}(b_j) = \sum_{i=1}^{m} \alpha_{i,j} c_i$. Let $v = \sum_j \gamma_j b_j$. Then, $\mathcal{M}(v) = \sum_{i=1}^{m} \tau_i c_i$ where*

$$\tau = \begin{bmatrix} \tau_1 \\ \vdots \\ \tau_m \end{bmatrix} = \begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & \cdots & \vdots \\ \alpha_{m.1} & \cdots & \alpha_{m,n} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = M\beta$$

The above result says that if we make the canonical (bijective) association of v with β and of $\mathcal{M}(v)$ with τ, then the action of the operator $\mathcal{M}$ corresponds to matrix-vector multiplication as $\tau = M\beta$. The matrix M clearly depends on the particular choice of basis made and is called the *matrix representation of $\mathcal{M}$ with* respect to these basis.

Let us explain this result another way.

Since $\mathbf{V}$ is of dimension n, it is isomorphic to $\mathbf{C}^n$. In other words (see *Notes on Vector Spaces, B-11*) , there exists a bijection $\phi : \mathbf{V} \to \mathbf{C}^n$ such that for all $v_1, v_2 \in \mathbf{V}$ and $\alpha \in \mathbf{C}$

$$\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2), \quad \text{and} \quad \phi(\alpha v_1) = \alpha\phi(v_1)$$

This isomorphism can be exhibited explicity as follows. Fix a basis $B = \{b_1, \cdots, b_n\}$ for $\mathbf{V}$. Any vector $v \in \mathbf{V}$ can be expressed as a (unique) linear combination $v = \sum_{i=1}^{n} \alpha_i b_i$. With the *abstract* vector $v \in \mathbf{V}$, we associate the *concrete* vector

$$\phi(v) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbf{C}^n$$

Of course the concrete representation depends on the choice of basis B.

In a similar fashion, given a basis $C = \{c_1, \cdots, c_n\}$ for $\mathbf{W}$, we can exhibit a linear bijection $\psi : \mathbf{W} \to \mathbf{C}^m$.

Now let $\mathcal{M} : \mathbf{V} \to \mathbf{W}$ be a linear operator. The main theorem in this point is that the action of $\mathcal{M}$ on an abstract vector v is equivalent to multiplication of a matrix $M \in \mathbf{C}^{m\times n}$ with the concrete representation $\phi(v)$ of the vector v. This is explained by saying that *the following diagram commutes:*

$$\begin{tikzcd}
\mathbf{V} \arrow[r, "\mathcal{M}"] \arrow[d, shift right, "\phi"'] & \mathbf{W} \arrow[d, shift right, "\psi"'] \\
\mathbf{C}^n \arrow[r, "M"'] \arrow[u, shift right, "\phi^{-1}"'] & \mathbf{C}^m \arrow[u, shift right, "\psi^{-1}"']
\end{tikzcd}$$

Armed with this association between matrices and linear operators on finite-dimensional vector spaces, we shall focus on matrix theory. Once our native intuition on matrices is sound, we will return to treat linear operators on infinite-dimensional vector spaces.

C. Change of Basis, Range Spaces, and Null Spaces

1 Change of bases.

Theorem *Consider a linear operator* $\mathcal{A} : \mathbf{V} \to \mathbf{W}$ *where* $dim(\mathbf{V}) = n, dim(\mathbf{W}) = m$. *Let*

$$B = \{b_1, \cdots, b_n\} \quad and \quad \hat{B} = \left\{\hat{b}_1, \cdots, \hat{b}_n\right\}$$

be two basis for $\mathbf{V}$ *and let*

$$\hat{B} = BT, \quad T \in \mathbf{C}^{n \times n}$$

Similarly let

$$C = \{c_1, \cdots, c_m\} \quad and \quad \hat{C} = \{\hat{c}_1, \cdots, \hat{c}_m\}$$

be two basis for $\mathbf{V}$ *and let*

$$\hat{C} = CR, \quad T \in \mathbf{C}^{m \times m}$$

Let A *and* $\hat{A}$ *be the matrix representations of the operator* ϕ *with respect to the basis* B, C *and* $\hat{B}, \hat{C}$ *respectively. Then,*

(a) R *and* T *are nonsingular.*

(b) $\hat{A} = R^{-1}AT$

A particularly important case of the above result is when $\mathbf{V} = \mathbf{W}$. Let $\hat{B} = BT$ relate the basis B and $\hat{B}$ of $\mathbf{V}$ and let A and $\hat{A}$ be the matrix representations of the operator ϕ with respect to the basis B and $\hat{B}$ respectively. Then,

(a) T is nonsingular.

(b) $\hat{A} = T^{-1}AT$

We therefore conclude that the square matrices A and $T^{-1}AT$ represent the same operator, albeit with respect to different basis. These matrices are said to be *similar* and the transformation T is called a *similarity transformation.* As one would expect, similar matrices have a lot in common, for instance, they have the same eigenvalues. They do not, however, have the same eigenvectors. Indeed, if v is an eigenvector of A with eigenvalue λ, then $T^{-1}v$ is an eigenvector of $\hat{A}$ corresponding to the same eigenvalue.

2 **Example** (Rotation operators)

Consider the vector space $\mathbf{R}^2$ and let ϕ be the operator that rotates a given vector by θ° counterclockwise. Then, the matrix representation of ϕ with respect to the standard basis is

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

3 **Definition** The *range space* $\mathcal{R}(T)$ of a linear operator T is the set

$$\mathcal{R}(T) = \{y \in \mathbf{W} : \ T(x) = y \text{ for some } \ x \in \mathbf{V}\}$$

The *null space* $\mathcal{N}(T)$ of T is the set

$$\mathcal{N}(T) = \{x \in \mathbf{V} : \ T(x) = 0\}$$

It is easy to verify that $\mathcal{R}(T)$ and $\mathcal{N}(T)$ are subspaces of $\mathbf{W}$ and $\mathbf{V}$ respectively.

Let $A \in \mathbf{C}^{m\times n}$ be a linear operator. Then $\mathcal{R}(A)$ is simply the set of all linear combinations of the columns of A and $\mathcal{N}(A)$ is the set of vectors $x \in \mathbf{C}^n$ such that $Ax = 0$.

4 **Lemma** *Let A, B be complex matrices of compatible dimensions.*

(a) $(A^*)^* = A$

(b) $(A + B)^* = A^* + B^*$

(c) $(AB)^* = B^*A^*$

(d) If A is nonsingular, then $(A^*)^{-1} = (A^{-1})^*$

5 We shall make extensive use of the following result.

Theorem *Let $A \in \mathbf{C}^{m\times n}$. Then*

(a) $\mathcal{R}^\perp(A) = \mathcal{N}(A^*)$

(b) $\mathbf{C}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^*)$

(c) $\mathcal{N}(A^*A) = \mathcal{N}(A)$

(d) $\mathcal{R}(AA^*) = \mathcal{R}(A)$

6 **Definition** The *rank* of a matrix A is the dimension of $\mathcal{R}(A)$.

The *nullity* of a matrix A is the dimension of $\mathcal{N}(A)$.

7 **Theorem** *Let $A \in \mathbf{C}^{m\times n}$ and $B \in \mathbf{C}^{n\times r}$. Then*

(a) $\textit{rank}\ (A) = \textit{rank}\ (A^*)$

(b) $\textit{rank}\ (A) \leq \min\{m, n\}$

(c) $\textit{rank}\ (A) + \textit{nullity}\ (A^*) = m$

(d) $\textit{rank}\ (A^*) + \textit{nullity}\ (A) = n$

(e) $\textit{rank}\ (A) = \textit{rank}\ (AA^*) = \textit{rank}\ (A^*A)$

(f) (Sylvester's inequality)

$$\textit{rank}\ (A) + \textit{rank}\ (B) - n \leq \textit{rank}\ (AB) \leq \min\{\textit{rank}\ (A), \textit{rank}\ (B)\}$$

8 **Theorem** *Let* $A \in \mathbf{C}^{m \times n}$ *and* $C \in \mathbf{C}^{n \times n}$ *and suppose* C *is invertible. Then,* $\mathcal{R}(A) = \mathcal{R}(AC)$ *and thus* $\textit{rank}\ (A) = \textit{rank}\ (AC)$

D. Eigenvectors and Eigenvalues

1 **Definition** Let $A \in \mathbf{C}^{n\times n}$. The *characteristic polynomial* of A written

$$\chi(s) = det(sI - A) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s^1 + \alpha_0 s^0$$

The roots of $\chi(s)$ of which there are n are called the *eigenvalues* of A.

Lemma *The eigenvalues of $A \in \mathbf{C}^{n\times n}$ are continuous functions of the entries of A.*

Lemma *Let $A \in \mathbf{C}^{n\times n}$ and λ be an eigenvalue of A. Then, there exists a vector $v \neq 0$ such that $Av = \lambda v$. This vector v is called an* eigenvector *of A with associated eigenvalue λ.*

It is evident that any nonzero multiple of v is also an eigenvector corresponding to the same eigenvalue. We view an eigenvector as a vector whose direction is invariant under the action of A. In the context of this interpretation, we regard all (nonzero) scalar multiples of an eigenvector as being the *same* eigenvector.

Let A have distinct eigenvalues. Then it has a *full* set of (i.e. n linearly independent) eigenvectors. This assertion will be proved shortly. If however A has *repeated eigenvalues a pathology may arise in that we are unable to find a full set of eigenvectors.*

Definition A matrix $A \in \mathbf{C}^{n\times n}$ is called *simple* if it has distinct eigenvalues. A matrix A is called *semi-simple* if it has n linearly independent eigenvectors.

2 **Definition** Let $A \in \mathbf{C}^{n\times n}$. A *subspace* $\mathcal{S} \subseteq \mathbf{C}^n$ is called A*-invariant* if $Ax \in \mathcal{S}$ for all $x \in \mathcal{S}$.

Theorem *Let $A \in \mathbf{C}^{n\times n}$ and let $\mathcal{S} \subseteq \mathbf{C}^n$ be an A-invariant subspace. Then $\mathcal{S}$ contains atleast one eigenvector of A.*

Theorem *Let $A \in \mathbf{C}^{n\times n}$ and let $\mathcal{S} \subseteq \mathbf{C}^n$ be an A-invariant subspace. Then $\mathcal{S}^\perp$ is A^*-invariant.*

Using these results, one can prove the following

Theorem *Let $A, B \in \mathbf{C}^{n\times n}$ be commuting matrices, i.e. $AB = BA$. Then, A and B share a common eigenvector.*

3 Simple-case of the Jordan form

Let $\mathcal{A} : \mathbf{C}^n \to \mathbf{C}^n$ be a linear operator and let A be the matrix representation of this operator with respect to some basis, say the standard basis of $\mathbf{C}^n$. We wish to make a special choice of basis so that the (new) matrix representation of $\mathcal{A}$ has particularly lucid structure. Equivalently, we wish to select a nonsingular matrix T in order that $T^{-1}AT$ has a transparent form called the *Jordan canonical from.*

In the special case where A is simple, we have the following key result.

Theorem *Let $A \in \mathbf{C}^{n\times n}$ be simple with eigenvalues λ_i and corresponding eigenvectors v_i, for $i = 1, \cdots, n$. Then*

(a) The eigenvectors form a basis for $\mathbf{C}^n$.

(b) Define the (nonsingular) matrix T and the matrix Λ by

$$T = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \quad , \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Then,

$$T^{-1}AT = \Lambda$$

The matrix Λ above is called the *Jordan form* of A. Observe that this matrix is diagonal and its diagonal entries are the eigenvalues of A.

4 **Example** Let

$$A = \begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix}$$

This matrix is simple and has eigenvalues at $1, -1$ with associated eigenvectors $\begin{bmatrix} 3 & 1 \end{bmatrix}'$ and $\begin{bmatrix} 1 & 1 \end{bmatrix}'$. It then follows that

$$T^{-1}AT = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \text{where} \quad T = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$$

5 General case of the Jordan form

For an arbitrary matrix $A \in \mathbf{C}^{n\times n}$ the situation is significantly more complex. Let $\lambda_1, \cdots, \lambda_k$ be the eigenvalues of A with associated multiplicities $m_1, \cdots, m_k$. Observe that $\sum_i m_i = n$.

We have the following result:

Theorem *There exists a nonsingular matrix $T \in \mathbf{C}^{n\times n}$ such that*

$$T^{-1}AT = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & J_k \end{bmatrix}$$

where the Jordan block $J_\ell \in \mathbf{C}^{\ell \times \ell}$ *corresponding to the eigenvalue* λ_ℓ *has the structure*

$$J_\ell = \begin{bmatrix} J_{1,\ell} & 0 & \cdots & 0 \\ 0 & J_{2,\ell} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & J_{r_\ell,\ell} \end{bmatrix}$$

and the Jordan sub-blocks *are as*

$$J_{i,\ell} = \begin{bmatrix} \lambda_\ell & 1 & \cdots & 0 & 0 \\ 0 & \lambda_\ell & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_\ell & 1 \\ 0 & 0 & \cdots & 0 & \lambda_\ell \end{bmatrix}$$

In order to determine the Jordan form of A we require the sizes and number of the various sub-blocks $J_{i,\ell}$ as well as the associated transformation matrix T. This is a complicated task, and we shall not delve into it. We should point out the following:

(a) The eigenvalues of A, including multiplicity, appear on the diagonal of the Jordan form.

(b) If A is semi-simple, its Jordan form is diagonal.

(c) In the general case, the Jordan form may have $1's$ along portions of the super-diagonal.

(d) The Jordan form of a matrix A is *not* a *continuous* function of the entries of A. To see this consider

Example

$$A_\epsilon = \begin{bmatrix} 1+\epsilon & 1 \\ 0 & 1 \end{bmatrix}$$

The Jordan form of A_ϵ for $\epsilon \neq 0$ is

$$\begin{bmatrix} 1+\epsilon & 0 \\ 0 & 1 \end{bmatrix}$$

while the Jordan form for A_0 is

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

As a result one should steer clear of algorithms that require *computation* of the Jordan form of a matrix.

E. Matrix polynomials and functions

1 Let $A \in \mathbf{C}^{n\times n}$ and let $p(s) = \sum_{i=0}^{k} \alpha_i s^i$ be a polynomial. Then, we *define*

$$p(A) = \sum_{i=0}^{k} \alpha_i A^i \in \mathbf{C}^{n\times n}$$

where $A^0 = I$.

We can generalize this notion to arbitrary (analytic) functions as follows. Consider the Taylor series

$$f(s) = \sum_{i=0}^{\infty} \alpha_i s^i$$

and *assume* that this Taylor series converges on $Spec(A)$. Then, we *define*

$$f(A) = \sum_{i=0}^{\infty} \alpha_i A^i \in \mathbf{C}^{n\times n}$$

(it will happen that this defining Taylor series also converges).

Lemma *Let $f(s), g(s)$ be arbitrary functions and let $h(s) = f(s)g(s)$. Then*

(a) $f(A)g(A) = g(A)f(A) = h(A)$

(b) $f(T^{-1}AT) = T^{-1}f(A)T$

(c) $$f\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\right) = \begin{bmatrix} f(A) & 0 \\ 0 & f(B) \end{bmatrix}$$

2 Computing functions of a matrix

We begin with the case where $A \in \mathbf{C}^{n\times n}$ is simple. In this case, the Jordan form $J = T^{-1}AT$ of A is diagonal and may be readily computed as in item (D-3). We can then employ properties (b) and (c) above to write

$$f(A) = Tf(J)T^{-1} = T \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{bmatrix} T^{-1}$$

Observe that $f(\lambda_i)$ are well-defined because $f(s)$ converges on $Spec(A)$. The reason we impose this requirement in defining $f(A)$ is now transparent.

Example Let A be as in item (D-4). We compute A^{300}:

$$A^{300} = TJ^{300}T^{-1} = T\begin{bmatrix} 1^{300} & 0 \\ 0 & -1^{300} \end{bmatrix} T^{-1} = I$$

We now turn our attention to the general case. Again, we shall proceed via the Jordan canonical form (see item (D-5)). Using the same idea, we see that we need to be able to compute $f(J)$ for a general Jordan block

$$J = \begin{bmatrix} \lambda & 1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

It turns out that (try to prove this !!!)

$$f(J) = \begin{bmatrix} f(\lambda) & \frac{1}{1!}f^{(1)}(\lambda) & \cdots & \frac{1}{(k-2)!}f^{(k-2)}(\lambda) & \frac{1}{(k-1)!}f^{(k-1)}(\lambda) \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & f(\lambda) & \frac{1}{1!}f^{(1)}(\lambda) \\ 0 & 0 & \cdots & 0 & f(\lambda) \end{bmatrix}$$

Observe that the derivatives $f^{(i)}(\lambda)$ above exist. Why? We are therefore in a position to compute (analytic) functions of an arbitrary matrix.

Example Consider the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

Using the above result, we obtain

$$\cos(A) = \begin{bmatrix} \cos(2) & -\sin(2) \\ 0 & \cos(2) \end{bmatrix}$$

3 The Spectral Mapping Theorem.

Theorem *Let $A \in \mathbf{C}^{n\times n}$ and let $f(s)$ be an arbitrary analytic function.*

(a) *Suppose A has eigenvalues $\{\lambda_1, \cdots, \lambda_n\}$. Then, the eigenvalues of $f(A)$ are $\{f(\lambda_1), \cdots, f(\lambda_n)\}$.*

(b) *Let v be an eigenvector of A with associated eigenvalue λ. Then v is also an eigenvector of $f(A)$ with associated eigenvalue $f(\lambda)$.*

4 Matrix exponentials.

Matrix exponentials are particularly important and arise in connection with systems of coupled linear ordinary differential equations. Since the Taylor series

$$e^{st} = 1 + st + \frac{s^2t^2}{2!} + \frac{s^3t^3}{3!} + \cdots$$

converges everywhere, we can define the exponential of any matrix $A \in \mathbf{C}^{n\times n}$ by

$$e^{At} = I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \cdots$$

5 Properties of Matrix exponentials.

Theorem

(a) $$e^{A0} = I$$

(b) $$e^{A(t+s)} = e^{At}e^{As}$$

(c) *If $AB = BA$ then*

$$e^{(A+B)t} = e^{At}e^{Bt} = e^{Bt}e^{At}$$

(d) $$det \left[\ e^{At}\ \right] = e^{traceAt}$$

(e) *e^{At} is nonsingular for all $-\infty < t < \infty$ and*

$$\left[\ e^{At}\ \right]^{-1} = e^{-At}$$

(f) *e^{At} is the* unique *solution X of the linear system of ordinary differential equations*

$$\dot{X} = AX, \quad \textit{subject to} \quad X(0) = I$$

6 Computing Matrix exponentials.

We may compute matrix exponentials via the method outlined in item (2).

Example Let A be as in item (D-4). We compute e^{At}:

$$e^{At} = Te^{Jt}T^{-1} = T\begin{bmatrix} e^t & 0 \\ 0 & e^{-t} \end{bmatrix} = \begin{bmatrix} 1.5e^t - 0.5e^{-t} & -1.5e^t + 1.5e^{-t} \\ 0.5e^t - 0.5e^{-t} & -0.5e^t + 1.5e^{-t} \end{bmatrix}$$

Example Consider the matrix $J \in \mathbf{C}^{n\times n}$ as

$$J = \begin{bmatrix} \lambda & 1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

This matrix is a typical component in the Jordan form of an arbitrary matrix. We shall compute e^{Jt}. We may, of course, employ the expression given in item (2), but we shall take a different route. First observe that

$$J = \lambda I + \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} = A + N$$

and that $AN = NA$. Thus $e^{Jt} = e^{At}e^{Nt}$. Since N is nilpotent with index n, we have

$$e^{Nt} = I + Nt + \frac{N^2t^2}{2!} + \cdots + \frac{N^{n-1}t^{n-1}}{(n-1)!} = \begin{bmatrix} 1 & t & \frac{1}{2!}t^2 & \cdots & \frac{1}{(n-1)!}t^{(n-1)} \\ 0 & 1 & t & \cdots & \frac{1}{(n-2)!}t^{(n-2)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

from which we immediately have e^{At}.

As an alternate method for computing matrix exponentials, we have the following

Theorem

$$e^{At} = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\}$$

where $\mathcal{L}$ denotes the Laplace transformation.

Example Consider the matrix

$$A = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$$

Using the above result we obtain

$$\begin{aligned} e^{At} &= \mathcal{L}^{-1}\left\{\frac{1}{(s-\sigma)^2+\omega^2}\begin{bmatrix} s-\sigma & \omega \\ -\omega & s-\sigma \end{bmatrix}\right\} \\ &= e^{\sigma t}\begin{bmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{bmatrix} \end{aligned}$$

7 **The Cayley-Hamilton Theorem.**

We shall make frequent use of the following fundamental result.

Theorem *Let* $A \in \mathbf{C}^{n\times n}$ *and let*

$$\chi(s) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s^1 + \alpha_0$$

be the characteristic polynomial of A*. Then,*

$$\chi(A) = 0$$

As a consequence of this theorem, it is evident that A^n is a linear combination of the set of matrices $\mathcal{S} = \{I, A, \cdots, A^{n-1}\}$. By an inductive argument it follows that any power of A, and therefore *any function* of A is expressible as a linear combination in $\mathcal{S}$.

In particular, we have

Lemma *Let* $A \in \mathbf{C}^{n\times n}$ *and let* $f(s)$ *be any (analytic) function for which* $f(A)$ *is defined. Then*

$$f(A) \in \mathcal{R}\left(\begin{bmatrix} I & A & \cdots & A^{(n-1)} \end{bmatrix}\right)$$

8 **Definition** Let $A \in \mathbf{C}^{n\times n}$. A polynomial $p(s)$ is called an *annihilating* polynomial of A if $p(A) = 0$. The *minimal* polynomial of A is the least degree, monic annihilating polynomial.

Lemma *Every annihilating polynomial of* A *is divisible by the minimal polynomial* $m(s)$ *of* A*.*

F. Hermitian and Definite Matrices

1 **Definition** A matrix $U \in \mathbf{C}^{n\times n}$ is called *unitary* if $U^*U = I = UU^*$.

A real unitary matrix is called an *orthogonal* matrix.

Lemma *Let $U \in \mathbf{C}^{n\times n}$ be unitary and consider the Hilbert space $\mathbf{C}^n$ equipped with the usual inner product. Then,*

(a) The columns of U form an orthonormal basis of $\mathbf{C}^n$.

(b) $\|Ux\| = \|x\|$

(c) $<Ux, Uy> = <x, y>$

(d) $U^{-1} = U^*$.

Rotation matrices (see item (B-2)) are unitary.

2 **Definition** A matrix $H \in \mathbf{C}^{n\times n}$ is called *Hermitian* if $H = H^*$.

Symmetric matrices are in particular Hermitian.

We will now prove several results regarding Hermitian matrices. These results also hold almost *verbatim* for symmetric matrices.

3 **Theorem** *The eigenvalues of a Hermitian matrix H are all real.*

4 **Theorem** *A Hermitian matrix H has a full set of eigenvectors. Moreover, these eigenvectors form an orthogonal set. As a consequence, Hermitian matrices can be diagonalized by unitary transformations, i.e. there exists a unitary matrix U such that*

$$H = UDU^*$$

where D is a diagonal matrix whose entries are the (real) eigenvalues of H.

5 **Theorem** *Let $H \in \mathbf{C}^{n\times n}$ be Hermitian. Then,*

(a) $$\sup_{v\neq 0} \frac{v^*Pv}{v^*v} = \lambda_{max}(P)$$

(b) $$\inf_{v\neq 0} \frac{v^*Pv}{v^*v} = \lambda_{min}(P)$$

6 **Definition** A matrix $P \in \mathbf{C}^{n\times n}$ is called *positive-definite* written $P > 0$ if P is Hermitian and further,

$$v^*Pv > 0\ , \quad \text{for all} \quad 0 \neq v \in \mathbf{C}^n$$

A matrix $P \in \mathbf{C}^{n \times n}$ is called *positive-semi-definite* written $P \geq 0$ if P is Hermitian and further,

$$v^* P v \geq 0 \ , \quad \text{for all} \ \ v \in \mathbf{C}^n$$

Analogous are the notions of *negative-* and *negative-semi-* definite matrices.

7 **Theorem** *Let $P \in \mathbf{C}^{n \times n}$ be Hermitian. The following are equivalent.*

(a) $P > 0$

(b) All the eigenvalues of P are positive.

(c) All the principal minors of P are positive.

Theorem *Let $P \in \mathbf{C}^{n \times n}$ be Hermitian. The following are equivalent.*

(a) $P \geq 0$

(b) All the eigenvalues of P are ≥ 0.

A principal minor test for positive-semi-definiteness is significantly more complicated.

8 **Lemma** *Let $0 < P \in \mathbf{C}^{n \times n}$ and let $X \in \mathbf{C}^{n \times m}$.*

(a) $\|x\|^2 = x^* P x$ qualifies as a norm on $\mathbf{C}^n$

(b) $X^* P X \geq 0$

(c) $X^* P X > 0$ if and only if rank $(X) = m$

9 **Definition** Let $0 \leq P \in \mathbf{C}^{n \times n}$. We can then write $P = UDU^*$ where U is unitary. Define the *square-root* of P written $P^{\frac{1}{2}}$ by

$$P^{\frac{1}{2}} = U D^{\frac{1}{2}} U^*$$

It is evident that $P^{\frac{1}{2}}$ as defined above is Hermitian, and moreover $P^{\frac{1}{2}} \geq 0$. Further, if $P > 0$, then $P^{\frac{1}{2}} > 0$.

G. The Singular-Value Decomposition

1 **Theorem** *Let $M \in \mathbf{C}^{m\times n}$ with rank $(M) = r$. Then we can find unitary matrices $U \in \mathbf{C}^{m\times m}$ and $V \in \mathbf{C}^{n\times n}$ such that*

$$M = U\Sigma V^* = U \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V^*$$

where

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}$$

The real numbers $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are called the singular values *of M and the representation above is called the* singular-value decomposition *of M.*

2 **Theorem** *Let $M \in \mathbf{C}^{m\times n}$ with rank $(M) = r$ and let $M = U\Sigma V^*$ be the singular-value decomposition of M. Partition U and V as*

$$U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}, \quad V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$$

where U_1 and V_1 are in $\mathbf{C}^{r\times r}$. Then

(a) The columns of U_1 and U_2 form orthonormal bases for $\mathcal{R}(M)$ and $\mathcal{N}(M^*)$ respectively.

(b) The columns of V_1 and V_2 form orthonormal bases for $\mathcal{R}(M^*)$ and $\mathcal{N}(M)$ respectively.

3 Computing the singular-value decomposition

While definitely not the method of choice *vis-a-vis* numerical aspects, the following result provides an adequate method for determining the singular-value decomposition of a matrix.

Theorem *Let $M \in \mathbf{C}^{m\times n}$ with rank $(M) = r$. Let $\lambda_1, \cdots, \lambda_r$ be the nonzero eigenvalues of M^*M. These will be nonnegative because $M^*M \geq 0$. Also, from item (F-4) it follows that there exist unitary matrices $U \in \mathbf{C}^{m\times m}$ and $V \in \mathbf{C}^{n\times n}$ such that*

$$MM^* = U \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} U^* \quad \text{and} \quad M^*M = V \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} V^*$$

where $\Lambda = diag\ (\lambda_1, \cdots, \lambda_r)$.

Then the singular-value decomposition of M is

$$M = U \begin{bmatrix} \Lambda^{\frac{1}{2}} & 0 \\ 0 & 0 \end{bmatrix} V^*$$

and the singular values of M are $\lambda_1^{\frac{1}{2}}, \cdots, \lambda_r^{\frac{1}{2}}$.

4 Optimal rank q approximations of matrices

Theorem *Let $M \in \mathbf{C}^{m\times n}$ with rank $(M) = r$ have singulat value decomposition as*

$$M = U \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V^*, \quad where \quad \Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}$$

Define the matrix

$$\hat{M} = U \begin{bmatrix} \hat{\Sigma}_1 & 0 \\ 0 & 0 \end{bmatrix} V^*, \quad where \quad \hat{\Sigma}_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma_q \end{bmatrix}$$

H. Linear Operators

1 We now return to treat linear operators on general (not necessarily finite dimensional) vector spaces. We focus on linear operators on normed spaces. Let $\mathbf{V}, \mathbf{W}$ be normed spaces.

Definition A linear operator

$$T : \mathbf{V} \to \mathbf{W}$$

is called *bounded* if

$$\sup_{v \in \mathbf{V}} \frac{\|T(v)\|_{\mathbf{W}}}{\|v\|_{\mathbf{V}}} < \infty$$

The set of $B(\mathbf{V}, \mathbf{W})$ bounded linear operators $T : \mathbf{V} \to \mathbf{W}$ forms a vector space under operation addition and operator scaling:

$$(T_1 + T_2)\, v = T_1(v) + T_2(v), \quad (\alpha T_1)\, v = \alpha(T_1 v)$$

It can readily be shown that $B(\mathbf{V}, \mathbf{W})$ is a *normed* space with the norm defined as

$$\|T\| = \sup_{v \in \mathbf{V}} \frac{\|T(v)\|_{\mathbf{W}}}{\|v\|_{\mathbf{V}}}$$

This definition of norm is called the *induced operator norm* as it is induced from the norms on $\mathbf{V}$ and $\mathbf{W}$.

2 **Examples**

3 **Theorem** *Let $\mathbf{V}, \mathbf{W}$ be normed spaces and consider the linear operator*

$$T : \mathbf{V} \to \mathbf{W}$$

Then, the following are equivalent:

(a) T is continuous everywhere.

(b) T is continuous at 0.

(c) T is bounded.

(d) $\mathcal{N}(T)$ is complete.

4 The situation becomes even more interesting when we consider linear operators on inner product spaces. We shall first require the following result.

Theorem *(Reisz Representation Theorem)*

Let $\mathbf{H}$ *be a Hilbert space and consider a bounded linear operator*

$$\phi : \mathbf{H} \to \mathbf{C}$$

Then, there exists a vector $x \in \mathbf{H}$ *such that for all* $h \in \mathbf{H}$,

$$\phi(h) = <x, h>$$

5 Let $\mathbf{V}$ and $\mathbf{W}$ be Hilbert spaces and consider a linear operator

$$T : \mathbf{V} \longrightarrow \mathbf{W}$$

Lemma *Fix* $w \in \mathbf{W}$. *Then there is a unique vector* $x \in \mathbf{V}$ *such that*

$$<w, v> = <x, v> \quad \textit{for all} \;\; v \in \mathbf{V}$$

This lemma allows us make the following

Definition The *adjoint* of T written T^* is the operator

$$T : \mathbf{W} \longrightarrow \mathbf{V}$$

defined (uniquely) by

$$<w, T(v)> = <T^*(w), v> \quad \text{for all} \;\; v \in \mathbf{V}$$

6 **Examples**